A class of asymptotic preserving schemes for kinetic equations and related problems with stiff sources

Francis Filbet, Shi Jin

A CLASS OF ASYMPTOTIC PRESERVING SCHEMES FOR KINETIC EQUATIONS AND RELATED PROBLEMS WITH STIFF SOURCES

FRANCIS FILBET AND SHI JIN

Abstract. In this paper, we propose a general framework to design asymptotic preserving schemes for the Boltzmann kinetic kinetic and related equations. Numerically solving these equations are challenging due to the nonlinear stiff collision (source) terms induced by small mean free or relaxation time. We propose to penalize the nonlinear collision term by a BGK-type relaxation term, which can be solved explicitly even if discretized implicitly in time. Moreover, the BGK-type relaxation operator helps to drive the density distribution toward the local Maxwellian, thus natually imposes an asymptotic-preserving scheme in the Euler limit. The scheme so designed does not need any nonlinear iterative solver or the use of Wild Sum. It is uniformly stable in terms of the (possibly small) Knudsen number, and can capture the macroscopic fluid dynamic (Euler) limit even if the small scale determined by the Knudsen number is not numerically resolved. It is also consistent to the compressible Navier-Stokes equations if the viscosity and heat conductivity are numerically resolved. The method is applicable to many other related problems, such as hyperbolic systems with stiff relaxation, and high order parabilic equations.

Contents

1. Introduction

The Boltzmann equation describes the time evolution of the density distribution of a dilute gas of particles when the only interactions taken into account are binary elastic collisions. For space variable $x \in \Omega \in \mathbb{R}^{d_x}$, particle velocity $v \in \mathbb{R}^{d_v}$ ($d_v \geq 2$), the Boltzmann eqaution reads:

$$\frac{\partial f}{\partial t} + v \cdot \nabla_x f = \frac{1}{\varepsilon} \mathcal{Q}(f) \tag{1.1}$$

where $f := f(t,x,v)$ is the time-dependent particles distribution function in the phase space. The parameter $\varepsilon > 0$ is the dimensionless Knudsen number which is the ratio the mean free path over a typical length scale such as the size of the spatial domain, thus measures the rarefiedness of the gas. The Boltzmann collision operator $\mathcal{Q}$ is a quadratic operator,

$$\mathcal{Q}(f)(v) = \int_{\mathbb{R}^{d_v}} \int_{\mathbb{S}^{d_v-1}} B(|v-v_\star|, \cos\theta)\ (f'_\star f' - f_\star f)\ d\sigma\, dv_\star. \tag{1.2}$$

F. Filbet is partially supported by the french ANR project "Jeunes Chercheurs" *Méthodes Numériques pour les Équations Cinétiques* (MNEC). S. Jin was partially supported by NSF grant No. DMS-0608720, NSF FRG grant DMS-0757285, and a Van Vleck Distinguished Research Prize from University of Wisconsin-Madison.

We used the shorthanded notation $f = f(v)$, $f_\star = f(v_\star)$, $f' = f(v')$, $f'_\star = f(v'_\star)$. The velocities of the colliding pairs $(v, v_\star)$ and $(v', v'_\star)$ are related by

$$\begin{cases} v' = v - \dfrac{1}{2}\big((v - v_\star) - |v - v_\star|\,\sigma\big), \\ \\ v'_\star = v - \dfrac{1}{2}\big((v - v_\star) + |v - v_\star|\,\sigma\big), \end{cases}$$

with $\sigma \in \mathbb{S}^{d_v - 1}$. The collision kernel B is a non-negative function which by physical arguments of invariance only depends on $|v - v_\star|$ and $\cos\theta = u \cdot \sigma$ (where $u = (v - v_\star)/|v - v_\star|$ is the normalized relative velocity). In this work we assume that B is locally integrable, given by

$$B(|u|, \cos\theta) = C_\gamma\,|u|^\gamma,$$

for some $\gamma \in (0, 1]$ and a constant $C_\gamma > 0$.

Boltzmann's collision operator has the fundamental properties of conserving mass, momentum and energy: at the formal level

$$\int_{\mathbb{R}^{d_v}} \mathcal{Q}(f)\,\phi(v)\,dv = 0, \qquad \phi(v) = 1, v, |v|^2, \tag{1.3}$$

and it satisfies the well-known Boltzmann's H theorem

$$-\frac{d}{dt}\int_{\mathbb{R}^{d_v}} f \log f\,dv = -\int_{\mathbb{R}^{d_v}} \mathcal{Q}(f)\log(f)\,dv \geq 0.$$

The functional $-\int f \log f$ is the *entropy* of the solution. Boltzmann's H theorem implies that any equilibrium distribution function, *i.e.*, any function which is a maximum of the entropy, has the form of a local Maxwellian distribution

$$\mathcal{M}_{\rho,u,T}(v) = \frac{\rho}{(2\pi T)^{d_v/2}} \exp\left(-\frac{|u - v|^2}{2T}\right),$$

where ρ, u, T are the *density*, *macroscopic velocity* and *temperature* of the gas, defined by

$$\rho = \int_{\mathbb{R}^{d_v}} f(v)\,dv = \int_{\mathbb{R}^{d_v}} \mathcal{M}\rho, u, T(v), \quad u = \frac{1}{\rho}\int_{\mathbb{R}^{d_v}} v\,f(v)\,dv = \frac{1}{\rho}\int_{\mathbb{R}^{d_v}} v\,\mathcal{M}\rho, u, T(v)\,dv \tag{1.4}$$

$$T = \frac{1}{d_v\rho}\int_{\mathbb{R}^{d_v}} |u - v|^2\,f(v)\,dv = \frac{1}{d_v\rho}\int_{\mathbb{R}^{d_v}} |u - v|^2\,\mathcal{M}\rho, u, T(v)\,dv \tag{1.5}$$

Therefore, when the Knudsen number $\varepsilon > 0$ becomes very small, the macroscopic description, which describe the evolution of averaged quantities such as the density ρ, momentum $\rho\,u$ and temperature T of the gas, by fluid dynamics equations, namely, the compressible Euler or Navier-Stokes equations, become adequate. Mor specifically, *i.e.* as $\varepsilon \to 0$, the distribution function will converge to a local Maxwellian $\mathcal{M}$, and the system (1.2) becomes a closed system for the $2 + d_v$ moments. The conserved quantities satisfy the classical Euler equations of gas dynamics for a mono-atomic gas:

$$\begin{cases} \dfrac{\partial \rho}{\partial t} + \nabla_x \cdot \rho\,u = 0, \\ \\ \dfrac{\partial \rho\,u}{\partial t} + \nabla_x \cdot (\rho\,u \otimes u + p\,\mathrm{I}) = 0, \\ \\ \dfrac{\partial E}{\partial t} + \nabla_x \cdot ((E + p)\,u) = 0, \end{cases} \tag{1.6}$$

where E represents the total energy

$$E = \frac{1}{2}\,\rho\,u^2 + \frac{d_v}{2}\,\rho\,T,$$

and I is the identity matrix. These equations constitute a system of $2 + d_v$ equations in $3 + d_v$ unknowns. The pressure is related to the internal energy by the constitutive relation for a polytropic gas

$$p = (\gamma - 1)\left(E - \frac{1}{2}\,\rho\,|u|^2\right),$$

where the polytropic constant $\gamma = (d_v + 2)/d_v$ represents the ratio between specific heat at constant pressure and at constant volume, thus yielding $p = \rho T$. For small but non zero values of the Knudsen number ε, the evolution equation for the moments can be derived by the so-called Chapman-Enskog expansion [8], applied to the Boltzmann equation. This approach gives the Navier-Stokes equations as a second order approximation with respect to ε to the solution to the Boltzmann equation:

$$\left\{\begin{array}{l} \dfrac{\partial \rho_\varepsilon}{\partial t} + \nabla_x \cdot \rho_\varepsilon\, u_\varepsilon = 0, \\ \dfrac{\partial \rho_\varepsilon\, u_\varepsilon}{\partial t} + \nabla_x \cdot (\rho_\varepsilon\, u_\varepsilon \otimes u_\varepsilon + p_\varepsilon\, \mathrm{I}) = \varepsilon\, \nabla_x \cdot [\mu_\varepsilon\, \sigma(u_\varepsilon)], \\ \dfrac{\partial E_\varepsilon}{\partial t} + \nabla_x \cdot (E_\varepsilon + p_\varepsilon)\, u_\varepsilon) = \varepsilon\, \nabla_x \cdot (\mu_\varepsilon \sigma(u_\varepsilon)\, u + \kappa_\varepsilon\, \nabla_x T_\varepsilon)\,. \end{array}\right. \tag{1.7}$$

In these equations $\sigma(u)$ denotes the strain-rate tensor given by

$$\sigma(u) = \nabla_x u + (\nabla_x u)^T - \frac{2}{d_v}\, \nabla_x \cdot u\, \mathrm{I}$$

while the viscosity $\mu_\varepsilon = \mu(T_\varepsilon)$ and the thermal conductivity $\kappa_\varepsilon = \kappa(T_\varepsilon)$ are defined according the linearized Boltzmann operator with respect to the local Maxwellian [1].

The connection between kinetic and macroscopic fluid dynamics results from two properties of the collision operator:

(*i*) conservation properties and an entropy relation that imply that the equilibria are Maxwellian distributions for the zeroth order limit;
(*ii*) the derivative of $\mathcal{Q}(f)$ satisfies a formal Fredholm alternative with a kernel related to the conservation properties of (*i*).

Past progress on developing robust numerical schemes for kinetic equations that also work in the fluid regimes has been guided by the fluid dynamic limit, in the framework of *asymptotic-preserving* (AP) scheme. As summarized by Jin [31], a scheme for the kinetic equation is AP if

- it preserves the discrete analogy of the Chapman-Enskog expansion, namely, it is a suitable scheme for the kinetic equation, yet, when holding the mesh size and time step fixed and letting the Knudsen number go to zero, the scheme becomes a suitable scheme for the limiting Euler equations
- implicit collision terms can be implemented explicitly, or at least more efficiently than using the Newton type solvers for nonlinear algebraic systems.

To satisfy the first condition for AP, the scheme must be driven to the local Maxwellian when $\epsilon \to 0$. This can usually be achieved by a backward Euler or any L-stable ODE solvers for the collision term [32]. Such a scheme requires an implicit collision term to gaurantee a uniform stability in time. However, how to invert such an implicit, yet nonlocal and nonlinear, collision operator is a delicate numerical issue. Namely, it is hard to realize the second condition for AP schemes.

Comparing with a multiphysics domain decomposition type method [4, 15, 17, 29, 39, 43], the AP schemes avoid the coupling of physical equations of different scales where the coupling conditions are difficult to obtain, and interface locations hard to determine. The AP schemes are based on solving one equation– the kinetic equation, and they become robust macroscopic (fluid) solvers *automatically* when the Knudsen number goes to zero. An AP scheme implying a numerical convergence uniformly in the Knudsen number was proved by Golse-Jin-Levermore for linear transport equation in the diffusion regime [27]. This result can be extended to essentially all AP schemes, although the specific proof is problem dependent. For examples of AP schemes for kinetic equations in the fluid dynamic or diffusive regimes see for examples [12, 5, 35, 36, 34, 37, 38, 28, 2, 40]. The AP framework has also been extended in [13, 14] for the study of the quasi-neutral limit of Euler-Poisson and Vlasov-Poisson systems, and in [16, 30] for all-speed (Mach number) fluid equations bridging the passage from compressible flows to the incompressible flows.

Since the Boltzmann collision term $\mathcal{Q}$ needs to be treated implicitly, how to invert it numerically becomes a tricky issue. One solution was offered by Gabetta, Pareschi and Toscani [25]. They first penalize $\mathcal{Q}$ by a linear function λf, and then absorb the linearly stiff part into the time variable to remove the stiffness. The remaining implicit nonlinear collision term is approximated by finite terms

in the Wild Sum, with the infinite sum replaced by the local Maxwellian. This yields a uniformly stable AP scheme for the collision term, capruring the Euler limit when $\epsilon \to 0$. Such a time-relaxed method was also used to develop AP Monte Carlo method, see [6, 41]. Nevertheless, it seems that this method is not able to capture the compressible Navier-Stokes asymptotic for small ε.

When the collision operator $\mathcal{Q}$ is the BGK collision operator

$$\mathcal{Q}_{BGK} = \mathcal{M} - f\,, \tag{1.8}$$

it is well-known that even an implicit collision term can be solved explicitly, using the property that $\mathcal{Q}$ preserves mass, momentum and energy. *Our new idea* is this paper is to utilize this property, and penalize the Boltzmann collision operator $\mathcal{Q}$ by the BGK operator:

$$\mathcal{Q} = [\mathcal{Q} - \lambda(\mathcal{M} - f)] + \lambda[\mathcal{M} - f] \tag{1.9}$$

where λ is the largest spectrum of the linearized collision operator of $\mathcal{Q}$ around the local Maxwellian $\mathcal{M}$. Now the first term on the right hand side of (1.9) is either not stiff, or less stiff compared to the second term, thus can be discretized *explicitly*, so as to avoid inverting the nonlinear operator $\mathcal{Q}$. The second term on the right hand side of (1.9) is stiff, thus will be treated implicitly. Despite this, as mentioned earlier, the implicit BGK operator can be inverted explicitly. Therefore we arrive at a scheme which is uniformly stable in ϵ, with an implicit source term that can be solved explicitly. In other words, in terms of handling the stiffness, the general Boltzmann collision operator can be handled as easily as the much simpler BGK operator, thus we significantly simplies an implicit Boltzmann solver!

Although a linear penalty (by removing $\mathcal{M}$ on the right hand size from (1.9) can also remove the stiffness, it does not have the AP property, unless one follows the Wild Sum procedure of [25]. The BGK operator that we use in (1.9) helps to drive f into $\mathcal{M}$, thus preserves the Euler limit. This will be proved asymptotically for prepared initial data (namely data near $\mathcal{M}$), and demonstrated numerically even for general initial data. Moreover, we will prove asymptotically that, for suitably small time-step, this method is also consistent to the Navier-Stokes equations (1.7) for $\epsilon << 1$.

Our method is partly motivated by the work of Haack, Jin and Liu [30], where by subtracting the leading linear part of the pressure in the compresible Euler equations with a low Mach number, the nonlinear stiffness in the pressure term due to the low Mach number is removed and an AP scheme was proposed for the compressible Euler or Navier-Stokes equations that capture the incompressible Euler or Navier-Stokes limit when the Mach number goes to zero.

Our method is not restricted to the Boltzmann equation. It applies to general nonlinear hyperbolic systems with stiff nonlinear relaxation terms [10, 33, 32, 11], and higher-order parabolic equations (see section 5). Moreover, it applies to any *stiff source term that admits a stable local equilibrium.*

In the following sections, we present a class of asymptotic preserving schemes designed for kinetic equations even if the general framework can be applied to other partial differential equations. We will focus on the Boltzmann equation and its hydrodynamic limit. We present different numerical tests to illustrate the efficiency of the present method. We treat particularly a multi-scale problem where the Knudsen numer ε depends on the space variable and takes different values ranging from 10^{-4} (hydrodynamic regime) to one (kinetic regime). Finally, the last part is devoted to the design of numerical schemes for nonlinear Fokker-Planck equations for which the asymptotic preserving scheme can be used to remove the CFL constraint of a parabolic equation.

2. An Asymptotic Preserving (AP) stiff ODE solver

Since out method does not depend on the discretization of the spatial derivative, but only on the structure of the stiff source term, we will first present in in the simplest framework for stiff ordinary differential equations.

Let us consider a Hilbert space H and the following nonlinear autonomous ordinary differential system

$$\left\{ \begin{array}{l} \dfrac{df_\varepsilon}{dt}(t) \,=\, \dfrac{\mathcal{Q}(f_\varepsilon(t))}{\varepsilon}, \qquad t \,\geq\, 0, \\ \\ f_\varepsilon(0) \,=\, f_0 \,\in\, H, \end{array} \right. \tag{2.1}$$

where the source term $\mathcal{Q}(f)$ satisfies the following properties:

- there exists a unique stationary solution $\mathcal{M}$ to (2.1), which satisfies $\mathcal{Q}(\mathcal{M}) = 0$;

- the solution to (2.1) converges to the steady state $\mathcal{M}$ when time goes to infinity, and the spectrum of $\nabla\mathcal{Q}(f) \subset \mathbb{C}^- = z \in \mathbb{C}^-, \mathrm{Im}(z) < 0$,

$$0 < \alpha \leq \|\nabla\mathcal{Q}(f)\| \leq L, \quad \forall\, f \in H.$$

Remark 2.1. *The second hypothesis above is certainly not the most general, but is convenient for our purpose. The lower bound implies that the solution converges to the steady state $\mathcal{M}$, while the upper bound is a sufficient condition for existence and uniqueness of a global solution.*

When ε becomes small, the differential equation (2.1) becomes stiff and explicit schemes are subject to severe stability constraints. Of course, implicit schemes allow larger time step, but new difficulty arises in seeking the numerical solution of a fully nonlinear problem at each time step. Here we want to combine both advantages of implicit and explicit schemes : large time step for stiff problems and low computational cost of the numerical solution at each time step.

We denote by f^n an approximation of $f(t^n)$ with $t^n = n\,\Delta t$ and the time step $\Delta t > 0$, Two classical procedures handle the aforementioned difficulties well. One is to linearize the unknown $Q(f^{n+1})$ at time step t^{n+1} around f at the previous time step f^n:

$$\mathcal{Q}(f^{n+1}) \approx \mathcal{Q}(f^n) + \nabla\mathcal{Q}(f^n)(f^{n+1} - f^n)$$

yielding a problem that only needs to solve a linear system with coefficient matrix depending on $\nabla\mathcal{Q}(f^n)$ [44]. This approach gives a uniformly stable time discretization without nonlinear solvers, however, it is not AP since the right hand size, as $\epsilon \to 0$, does not project f^{n+1} to the local equilibrium $f = \mathcal{M}$, even if $f^n = \mathcal{M}$. The second approach, introduced by [25], takes

$$\mathcal{Q}(f) = [\mathcal{Q}(f) - \mu f] + \mu f\,.$$

As mentioned in the introduction, it uses the Wild Sum expansion for $\mathcal{Q}(f)$ on the right side, which is truncated and the remaining infinite series is replaced by the local Maxwellian in order to be AP for the Euler limit.

Under our hypothesis, the asymptotic behavior of the exact solution f_ε is known when $\varepsilon \to 0$. Therefore, we split the source term of (2.1) in a stiff and non- (or less) stiff part as

$$\frac{\mathcal{Q}(f)}{\varepsilon} = \underbrace{\frac{\mathcal{Q}(f) - P(f)}{\varepsilon}}_{\text{non stiff part}} + \underbrace{\frac{P(f)}{\varepsilon}}_{\text{stiff part}},$$

where $P(f)$ is a *well balanced*, *i.e.* preserving the steady state, $P(\mathcal{M}) = 0$, linear operator and is close to the source term $\mathcal{Q}(f)$. For instance, performing a simple Taylor expansion, we get

$$\mathcal{Q}(f) = \mathcal{Q}(\mathcal{M}) + \nabla\mathcal{Q}(\mathcal{M})\,(f - \mathcal{M}) + O(\|f - \mathcal{M}\|_H^2)$$

and we may choose

$$P(f) := \nabla\mathcal{Q}(\mathcal{M})\,(f - \mathcal{M}).$$

Since it is not always possible to compute exactly $\nabla\mathcal{Q}(\mathcal{M})$, we may simply choose

$$P(f) := L\,(f - \mathcal{M})\,,$$

where L is an estimate of $\nabla\mathcal{Q}(\mathcal{M})$.

Now, we simply apply a first order implicit-explicit (IMEX) scheme for the time discretization of (2.1):

$$\frac{f^{n+1} - f^n}{\Delta t} = \frac{\mathcal{Q}(f^n) - P(f^n)}{\varepsilon} + \frac{P(f^{n+1})}{\varepsilon}, \tag{2.2}$$

or

$$f^{n+1} = [\varepsilon\,\mathrm{I} - \Delta t\,\nabla\mathcal{Q}(\mathcal{M})]^{-1}\,[\varepsilon\,f^n + \Delta t\,(\mathcal{Q}(f^n) - P(f^n)) - \Delta t\,\nabla\mathcal{Q}(\mathcal{M})\,\mathcal{M}]\,.$$

This method is easy to implement, since f^{n+1} is linear in the right hand side of (2.2). For linear problems, we have the following result:

Theorem 2.2. *Consider the differential system (2.1) with $\mathcal{Q}(f) = -\lambda\,f$, where $Re(\lambda) > 0$. Set $P(f) := -\nu\,\lambda\,f$ with $\nu \geq 0$. Then, the scheme (2.2) is A-stable and L-stable for $\nu > 1/2$.*

Proof. For linear systems, the scheme simple reads

$$f^{n+1} \,=\, \frac{\varepsilon \,+\, (\nu-1)\,\lambda\,\Delta t}{\varepsilon \,+\, \nu\lambda\,\Delta t}\, f^n \,=\, \left(1 - \frac{\lambda\Delta t}{\varepsilon + \nu\,\lambda\Delta t}\right) f^n.$$

Observe that $\nu = 0$ gives the explicit Euler scheme, which is stable only for $\Delta t \le \varepsilon/\lambda$, whereas for $0 \le \nu \le 1$, it yields the so-called θ-scheme, which is A-stable for $\nu > 1/2$. For $\nu = 1$ it corresponds to the A-stable implicit Euler scheme. Moreover, for $\nu > 1$, the scheme is A-stable, that is

$$\|f^{n+1}\|_H \,\le\, \left|1 - \frac{\lambda\Delta t}{\varepsilon + \nu\,\lambda\Delta t}\right| \|f^n\|_H \sim \left(1 - \frac{1}{\nu}\right)\|f^n\|_H \quad \text{for} \quad \epsilon \sim 0\,,$$

where $|1 - \frac{1}{\nu}| < 1$ for $\nu > 1/2$. This is also the condition for the L-stability [26]. □

To improve the numerical accuracy, second order schemes are sometimes more desirable. Thus, we propose the following second order IMEX extension. Assume that an approximate solution f^n is known at time t^n, we compute a first approximation at time $t^{n+1/2} = t^n + \Delta t/2$ using a first order IMEX scheme and next apply the trapezoidal rule and the mid-point formula. The scheme reads

$$\left\{\begin{array}{l} 2\,\dfrac{f^\star - f^n}{\Delta t} \,=\, \dfrac{\mathcal{Q}(f^n) \,-\, P(f^n)}{\varepsilon} \,+\, \dfrac{P(f^\star)}{\varepsilon}. \\[2ex] \dfrac{f^{n+1} - f^n}{\Delta t} \,=\, \dfrac{\mathcal{Q}(f^\star) \,-\, P(f^\star)}{\varepsilon} \,+\, \dfrac{P(f^n) + P(f^{n+1})}{2\,\varepsilon}. \end{array}\right. \tag{2.3}$$

Note that both (2.2) and (2.3) are AP for prepared initial data. Let us use (2.3) as the example. Assume $f^n = \mathcal{M} + O(\varepsilon)$. Then $\mathcal{Q}(f^n) = O(\varepsilon), P(f^n) = O(\varepsilon)$, thus the first step in (2.3) gives $P(f^*) = O(\varepsilon)$. This further implies that $\mathcal{Q}(f^*) = O(\varepsilon)$. Applying all these in the second step of (2.3) gives $P(f^{n+1}) = O(\varepsilon)$, thus $f^{n+1} = \mathcal{M} + O(\varepsilon)$, which is the desired AP property.

To illustrate the efficiency of (2.2) and (2.3) in various situations, we consider a simple linear problem with different scales for which only some components rapidly converge to a steady state whereas the remaining part oscillates. We solve

$$\mathcal{Q}(f) \,=\, A\,f, \tag{2.4}$$

where

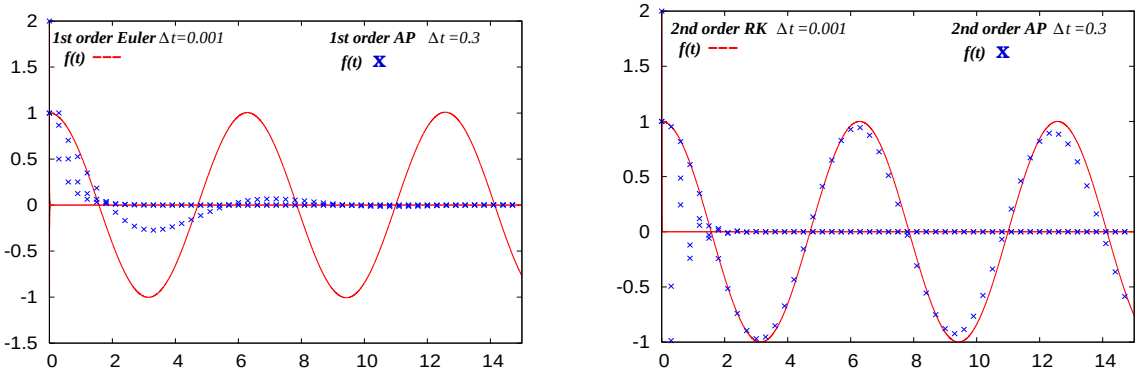
$$A = \left(\begin{array}{ccc} -1000 & 1 & 0 \\ -1 & -1000 & 0 \\ 0 & 0 & i \end{array}\right) \tag{2.5}$$

for which the eigenvalues are $\mathrm{Sp}(A) \,=\, \{-1000 + i, -1000 - i, i\}$. The first block represents the fast scales whereas the last one is the oscillating part. Indeed, the first components go to zero exponentially fast whereas the third one oscillates with respect to time with a period of 2π. We want to solve accurately the oscillating part with a large time step without resolving small scales. Then, we apply the first order (2.2) and second order (2.3) schemes by choosing

$$P(f) = \nu\,A\,f,$$

with $\nu \ge 0$. Here we take a large time step $\Delta t = 0.3$ and $\nu = 2$, which means that $P(f)$ has the same structure of $\mathcal{Q}(f)$ but the eigenvalues are over estimated. Thus, fast scales are under-resolved whereas this times step is a good discretization of the third oscillating component. Therefore, an efficient AP scheme would give an accurate behavior of the slow oscillating scale with large time step with respect to the fast scale. It clearly appears in Figure 1 that the time step is too large to give accurate results for the first order scheme (2.2): the solution is stable but the oscillation of the third component is damped for this time step which is too large. This approximation is compared with the one obtained with a first order explicit Euler using a times step ten times smaller. We also compare the numerical solution of the second order scheme (2.3) with the one obtained using a second order explicit Runge-Kutta scheme corresponding to $\nu = 0$ with a time step three hundred times smaller. In Figure 1, we observe the stability and good accuracy of the second order scheme (2.3). Note that for the same time step, the explicit Runge-Kutta scheme blows-up!

In the following sections we apply this approach to the Boltzmann equation and verify its accuracy and efficiency on several classical problems dealing with fluid, kinetic and multi-scale regimes.



(1) (2)

FIGURE 1. Comparison of (1) first and (2) second order Asymptotic Preserving and explicit Runge-Kutta schemes for the differential system (2.4)-(2.5).

3. Application to the Boltzmann equation

We now extend the stiff ODE solver of the previous section to the Boltzmann eqaution (1.1). To this aim, we rewrite the Boltzmann equation (1.1) in the following form

$$\left\{ \begin{array}{l} \dfrac{\partial f}{\partial t} + v\,\nabla_x f = \dfrac{\mathcal{Q}(f) - P(f)}{\varepsilon} + \dfrac{P(f)}{\varepsilon}, \quad x \in \Omega \subset \mathbb{R}^{d_x},\ v \in \mathbb{R}^{d_v}, \\ \\ f(0,x,v) = f_0(x,v), \quad x \in \Omega,\ v \in \mathbb{R}^{d_v}, \end{array} \right. \tag{3.1}$$

where the operator P is a "well balanced relaxation approximation" of $\mathcal{Q}(f)$, which means that it satisfies the following (balance law)

$$\int_{\mathbb{R}^d} P(f)\,\phi(v)\,dv = 0, \qquad \phi(v) = 1, v, |v|^2,$$

and preserves the steady state *i.e.* $P(\mathcal{M}_{\rho,u,T}) = 0$ where $\mathcal{M}_{\rho,u,T}$ is the Maxwellian distribution associated to ρ, u and T given by (1.4). Moreover, it is a relaxation operator in velocity

$$P(f) = \beta\,[\mathcal{M}_{\rho,u,T}(v) - f(v)]\,. \tag{3.2}$$

For instance, $P(f)$ can be computed from an expansion of the Boltzmann operator with respect to $\mathcal{M}\rho, u, T$:

$$\mathcal{Q}(f) \quad \simeq \quad \mathcal{Q}(\mathcal{M}_{\rho,u,T}) + \nabla\mathcal{Q}(\mathcal{M}_{\rho,u,T})\,[\mathcal{M}_{\rho,u,T} - f]\,.$$

Thus, we choose $\beta > 0$ as an upper bound of the operator $\nabla\mathcal{Q}(\mathcal{M}_{\rho,u,T})$. Then $P(f)$ given by (3.2) is just the BGK collisional operator [3].

Since the convection term in (3.1) is not stiff, we will treat it explicitly. For source terms on the right hand side of (3.1) will be handled using the ODE solver in the previous section. For example, if the first order scheme (2.2) is used, then we have

$$\left\{ \begin{array}{l} \dfrac{f^{n+1} - f^n}{\Delta t} + v \cdot \nabla_x f^n = \dfrac{\mathcal{Q}(f^n) - P(f^n)}{\varepsilon} + \dfrac{P(f^{n+1})}{\varepsilon}, \\ \\ f^0(x,v) = f_0(x,v)\,. \end{array} \right. \tag{3.3}$$

Using the relaxation structure of $P(f)$, it can be written as

$$\begin{aligned} f^{n+1} \quad = \quad & \frac{\varepsilon}{\varepsilon + \beta\Delta t}\,[f^n - \Delta t\, v\,\nabla_x f^n] + \Delta t\,\frac{\mathcal{Q}(f^n) - P(f^n)}{\varepsilon + \beta\Delta t} \\ & + \frac{\beta\Delta t}{\varepsilon + \beta\Delta t}\,\mathcal{M}^{n+1}, \end{aligned}$$

where $\mathcal{M}^{n+1}$ is the Maxwellian distribution computed from f^{n+1}.

Although (3.4) appears nonlinearly implicit, it can be computed explicitly. Specifically, upon multiplying (3.4) by $\phi(v)$ defined in (1.3), and use the conservation property of $\mathcal{Q}$ and P and the definition of $\mathcal{M}$ in (1.4), one gets

$$U^{n+1} = \frac{\varepsilon}{\varepsilon+\beta\Delta t}\int \phi(f^n - \Delta t\, v\cdot\nabla_x f^n)\,dv + \frac{\beta\Delta t}{\varepsilon+\beta\Delta t}U^{n+1}\,,$$

or simply

$$U^{n+1} = \int \phi(f^n - \Delta t v\cdot\nabla_x f^n)\,dv\,.$$

Thus U^{n+1} can be obtained explicitly, which defines $\mathcal{M}^{n+1}$. Now f^{n+1} can be obtained from (3.4 explicitly. In summary, although (3.3) is nonlinearly implicit, it can be solved *explicitly*, thus satisfies the second condition of an AP scheme.

We define the macroscopic quantity U by $U := (\rho, \rho\,u, T)$ computed from f. Clearly, the scheme (3.3) satisfies the following properties

Proposition 3.1. *Consider the numerical solution given by (3.3). Then,*

(i) If $\varepsilon \to 0$ and $f^n = \mathcal{M}^n + O(\varepsilon)$, then the scheme (3.3) is asymptotic preserving, that is $f^{n+1} = \mathcal{M}^{n+1} + O(\varepsilon)$, thus the scheme is AP to the Euler limit in the sence that, when $\varepsilon \to 0$, the (moments of the) scheme becomes a consistent discretization of the Euler system (1.6).

(ii) For $\varepsilon \ll 1$ and there exists a constant $C > 0$ such that

$$\left\|\frac{f^{n+1}-f^n}{\Delta t}\right\| + \left\|\frac{U^{n+1}-U^n}{\Delta t}\right\| \le C, \tag{3.4}$$

then the scheme (3.3) asymptotically becomes a first order in time approximation of the compressible Navier-Stokes (1.7).

Proof. We easily first check that for $\varepsilon \to 0$ and $f^n = \mathcal{M}^n$, we get $f^{n+1} = \mathcal{M}^{n+1}$. Therefore, we multiply (3.3) by $(1, v, |v|^2/2)$ and integrate with respect to v, which yields that U^n is given by a time explicit scheme of the Euler system (1.6).

Now let us prove (ii). We apply the classical Chapman-Enskog expansion:

$$f^n \;=\; \mathcal{M}^n \;+\; \varepsilon\, g^n \tag{3.5}$$

and integrate (3.3) with respect to $v \in \mathbb{R}^{d_v}$. By using the conservation properties of the Boltzmann operator (1.3) and of the well-balanced approximation $P(f)$,

$$\frac{U^{n+1}-U^n}{\Delta t} \;+\; \nabla_v \cdot \int_{\mathbb{R}^{d_v}} \begin{pmatrix} 1 \\ v \\ \dfrac{|v|^2}{2} \end{pmatrix} v\,(\mathcal{M}^n \;+\; \varepsilon g^n)dv \;=\; 0. \tag{3.6}$$

For $\varepsilon g = 0$, this is the compressible Euler equations (1.6). Thus, a consistent approximation of the compressible Navier-Stokes is directly related to a consistent approximation of g^n. Inserting decomposition (3.5) into the scheme (3.3) gives

$$\begin{aligned} \frac{\mathcal{M}^{n+1}-\mathcal{M}^n}{\Delta t} \;+\; & v\,\nabla_x\mathcal{M}^n \;+\; \varepsilon\left(\frac{g^{n+1}-g^n}{\Delta t} + v\,\nabla_x g^n\right) \\ & = \;\; \frac{\mathcal{Q}(\mathcal{M}^n+\varepsilon g^n)}{\varepsilon} \;-\; \left[\beta(\rho^n)g^n - \beta(\rho^{n+1})g^{n+1}\right], \end{aligned}$$

Since $\mathcal{Q}$ is a bilinear and $\mathcal{Q}(\mathcal{M}) = 0$, one has

$$\mathcal{Q}(\mathcal{M} + \varepsilon\, g) \;=\; \mathcal{Q}(\mathcal{M}) \;+\; \varepsilon\,\mathcal{L}_{\mathcal{M}}(g) + \varepsilon^2\,\mathcal{Q}(g),$$

where $\mathcal{L}_{\mathcal{M}}$ is the linearized collision operator with respect to $\mathcal{M}$. Thus, we get

$$\begin{aligned} & \frac{\mathcal{M}^{n+1}-\mathcal{M}^n}{\Delta t} \;+\; \left[\beta(\rho^n)g^n - \beta(\rho^{n+1})g^{n+1}\right] \\ +\;\; & \varepsilon\left[\frac{g^{n+1}-g^n}{\Delta t} + v\,\nabla_x g^n - \mathcal{Q}(g^n)\right] \\ & = \mathcal{L}_{\mathcal{M}}(g^n) \;-\; v\,\nabla_x\mathcal{M}^n\,, \end{aligned} \tag{3.7}$$

It is well known that $\mathcal{L}_{\mathcal{M}}$ is a non-positive self-adjoint operator on $L^2_{\mathcal{M}}$ defined by the set

$$L^2_{\mathcal{M}} := \{\varphi : \quad \varphi\,\mathcal{M}^{-1/2} \in L^2(\mathbb{R}^{d_v})\}$$

and that its kernel is $\mathcal{N}(\mathcal{L}_{\mathcal{M}}) = \mathrm{Span}\{\mathcal{M}, v\,\mathcal{M}, |v|^2\mathcal{M}\}$. Let $\Pi_{\mathcal{M}}$ be the orthogonal projection in $L^2_{\mathcal{M}}$ onto $\mathcal{N}(\mathcal{L}_{\mathcal{M}})$. After easy computations in the orthogonal basis, one finds that

$$\Pi_{\mathcal{M}}(\psi) = \frac{\mathcal{M}}{\rho}\left[m_0 + \frac{v-u}{T}m_1 + \left(\frac{|v-u|^2}{2T} - \frac{d_v}{2}\right)m_2\right]$$

where

$$m_0 = \int_{\mathbb{R}^{d_v}} \psi\,dv, \quad m_1 = \int_{\mathbb{R}^{d_v}} (v-u)\psi\,dv, \quad m_2 = \int_{\mathbb{R}^{d_v}} \left(\frac{|v-u|^2}{2T} - \frac{d_v}{2}\right)\psi\,dv.$$

It is easy to verify that $\Pi_{\mathcal{M}^n}(\mathcal{M}^n) = \mathcal{M}^n$ and

$$\Pi_{\mathcal{M}^n}(g^n) \;=\; \Pi_{\mathcal{M}^n}(g^{n+1}) \;=\; \Pi_{\mathcal{M}^n}(Q(g^n)) \;=\; \Pi_{\mathcal{M}^n}(\mathcal{L}_{\mathcal{M}^n}(g^n)) \;=\; 0.$$

Then applying the orthogonal projection $\mathrm{I} - \Pi_{\mathcal{M}^n}$ to (3.7), it yields

$$\begin{aligned}
&(\mathrm{I}-\Pi_{\mathcal{M}^n})\left(\frac{\mathcal{M}^{n+1}-\mathcal{M}^n}{\Delta t}\right) + \left(\beta(\rho^n)g^n - \beta(\rho^{n+1})g^{n+1}\right)\\
&+\varepsilon\left[\frac{g^{n+1}-g^n}{\Delta t} + (\mathrm{I}-\Pi_{\mathcal{M}^n})\,(v\,\nabla_x g^n) - \mathcal{Q}(g^n)\right]\\
=\;& \mathcal{L}_{\mathcal{M}}(g^n) - (\mathrm{I}-\Pi_{\mathcal{M}^n})\,(v\,\nabla_x \mathcal{M}^n).
\end{aligned}$$

Finally, it remains to estimate

$$(\mathrm{I}-\Pi_{\mathcal{M}^n})\left(\frac{\mathcal{M}^{n+1}-\mathcal{M}^n}{\Delta t}\right).$$

Using a Taylor expansion we find that

$$\mathcal{M}^{n+1} = \mathcal{M}^n\left[1 + \frac{\rho^{n+1}-\rho^n}{\rho^n} + \frac{v-u^n}{T^n}\left(u^{n+1}-u^n\right) + \left(\frac{|v-u^n|^2}{2T^n} - \frac{d}{2}\right)\frac{T^{n+1}-T^n}{T^n}\right] + O(\Delta t^2)$$

and by definition of $\Pi_{\mathcal{M}}$

$$\begin{aligned}
&\Pi_{\mathcal{M}^n}(\mathcal{M}^{n+1}) \;=\\
&\mathcal{M}^n\left(1 + \frac{\rho^{n+1}-\rho^n}{\rho^n} + \frac{v-u^n}{T^n}\left(u^{n+1}-u^n\right) + \left(\frac{|v-u^n|^2}{2T^n} - \frac{d}{2}\right)\frac{T^{n+1}-T^n}{T^n}\right]\\
+\;& \mathcal{M}^n\left(\frac{|v-u^n|^2}{2T^n} - \frac{d}{2}\right)\left[\frac{T^{n+1}-T^n}{\rho^n\,T^n}(\rho^{n+1}-\rho^n) + \frac{\rho^{n+1}}{d\rho^n\,T^n}(u^{n+1}-u^n)^2\right]\\
+\;& \mathcal{M}^n\,\frac{v-u^n}{T^n}\,\frac{\rho^{n+1}-\rho^n}{\rho^n}\left(u^{n+1}-u^n\right) + O(\Delta t^2).
\end{aligned}$$

Thus, under the assumption (3.4), we have

$$(\mathrm{I}-\Pi_{\mathcal{M}^n})\left(\frac{\mathcal{M}^{n+1}-\mathcal{M}^n}{\Delta t}\right) = O(\Delta t)$$

and the residual distribution function is given by

$$g^n \;=\; \mathcal{L}^{-1}_{\mathcal{M}^n}\big((\mathrm{I}-\Pi_{\mathcal{M}^n})\,(v\cdot\nabla_x\mathcal{M}^n)\big) + O(\varepsilon) \;+\; O(\Delta t).$$

Now, substituting this latter expression in (3.6), we get

$$\begin{aligned}
\frac{U^{n+1}-U^n}{\Delta t} + \nabla_x\cdot F(U) \;=\;& -\varepsilon\,\nabla_x\cdot\int_{\mathbb{R}^{d_v}}\begin{pmatrix} v\\ v\otimes v\\ v\dfrac{|v|^2}{2}\end{pmatrix}\mathcal{L}^{-1}_{\mathcal{M}^n}\big((\mathrm{Id}-\Pi_{\mathcal{M}^n})\,(v\cdot\nabla_x\mathcal{M}^n)\big)dv\\
&+O(\varepsilon\Delta t + \varepsilon^2),
\end{aligned}$$

where

$$F(U) = \begin{pmatrix} \rho\,u\\ \rho\,u\otimes u + p\,\mathrm{I}\\ (E\,+\,p)\,u\end{pmatrix}.$$

To complete the proof, it remains to compute the term in $O(\varepsilon)$. An easy computation first gives

$$(\mathrm{I} - \Pi_{\mathcal{M}^n})\,(v\cdot\nabla_x\mathcal{M}^n) \;=\; \left[B\,\left(\nabla u + (\nabla u)^T - \frac{d}{2}\nabla\cdot u\,\mathrm{I}\right) \;+\; A\,\frac{\nabla T}{\sqrt{T}}\right]\,M(v),$$

with

$$A = \left(\frac{|v-u|^2}{2T} - \frac{d+2}{2}\right)\frac{v-u}{\sqrt{T}}, \quad B = \frac{1}{2}\left(\frac{(v-u)\otimes(v-u)}{2T} - \frac{|v-u|^2}{dT}\mathrm{I}\right).$$

Therefore, it yields

$$\begin{aligned}\mathcal{L}^{-1}_{\mathcal{M}^n}\big(\,(\mathrm{I} - \Pi_{\mathcal{M}^n})\,(v\cdot\nabla_x\mathcal{M}^n)\,\big) \;&=\; \mathcal{L}^{-1}_{\mathcal{M}^n}(B\,M)\left(\nabla u + (\nabla u)^T - \frac{d}{2}\nabla\cdot u\,\mathrm{I}\right)\\ &\quad + \mathcal{L}^{-1}_{\mathcal{M}^n}(A\,M)\,\frac{\nabla T}{\sqrt{T}}.\end{aligned}$$

Substituting this expression in (3.6), we get a consistent time discretization scheme to the compressible Navier-Stokes system where the term of order of ε is given by

$$\varepsilon\,\nabla_x\cdot\begin{pmatrix}0\\ \mu_\varepsilon\,\sigma(u_\varepsilon)\\ \mu_\varepsilon\sigma(u_\varepsilon)\,u \;+\; \kappa_\varepsilon\,\nabla_x T_\varepsilon\end{pmatrix}$$

with

$$\sigma(u) \;=\; \nabla_x u \;+\; (\nabla_x u)^T \;-\; \frac{2}{d_v}\,\nabla_x\cdot u\,\mathrm{I}$$

while the viscosity $\mu_\varepsilon = \mu(T_\varepsilon)$ and the thermal conductivity $\kappa_\varepsilon = \kappa(T_\varepsilon)$ are defined according the linearized Boltzmann operator with respect to the local Maxwellian [1]. □

Remark 3.2. *To capture the Navier-Stokes approximation that has $O(\varepsilon)$ viscosity and heat conductivity, one needs the mesh size and $c\Delta t$ to be $o(\varepsilon)$ (c is a characteristic speed). Thus conclusion (ii) in the above proposition shows that the scheme is consistent to the Navier-Stokes equations provided that the viscous terms are resolved, while to capturing the Euler limit one can use mesh size and $c\Delta t$ much larger than ε, in the usual sense of asymptotic-preserving.*

4. Numerical tests

In this section we perform several numerical simulations for the Boltzmann equation in different asymptotic regimes in order to check the performance (in stability and accuracy) of our methods. We have implemented the first order (2.2) and second order (2.3) scheme for the approximation of the Boltzmann equation. Here, the Boltzmann collision operator is discretized by a deterministic method [18, 19, 20, 22], which gives a spectrally accurate approximation. A classical second order finite volume scheme with slope limiters is applied for the transport operator.

4.1. Approximation of smooth solutions.

This test is used to evaluate the order of accuracy of our new methods. More precisely, we want to show that our methods (2.2) and (2.3) are uniformly accurate with respect to the parameter $\varepsilon > 0$. We consider the Boltzmann equation (1.1) in $1\,d_x\times 2\,d_v$. We take a smooth initial data

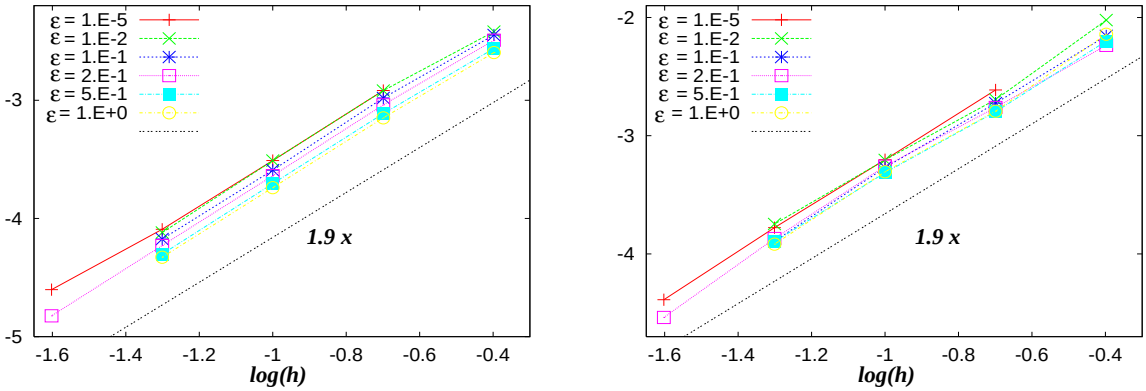
$$f_0(x,v) \;=\; \frac{\rho_0(x)}{2\pi\,T_0(x)}\,\exp\left(-\frac{|v|^2}{2\,T_0(x)}\right), \quad (x,v)\in[-L,L]\times\mathbb{R}^2,$$

with $\rho_0(x) = (11 - 9\tanh(x))/10$, $T_0(x) = (3-\tanh(x))/4$, $L = 1$ and assume specular reflection boundary conditions in x. Numerical solutions are computed from different phase space meshes : the number of point in space is $n_x = 50, 100, 200,\ldots,1600$ and the number of points in velocity is n_v^2 with $n_v = 8,\ldots,64$ (for which the spectral accuracy is achieved), the time step is computed such that the CFL condition for the transport is satisfied $\Delta t \le \Delta x/v_{\max}$, where Δx is the space step and $v_{\max} = 7$ is the truncation of the velocity domain. Then different values of ε are considered starting from the fully kinetic regime $\varepsilon = 1$, up to the fluid limit $\varepsilon = 10^{-5}$ corresponding to the solution of the Euler system (1.6). The final time is $T_{\max} = 1$ such that the solution is smooth for the different regimes.

An estimation of the relative error in L^p norm is given by

$$\varepsilon_{2\,h} = \max_{t\in(0,T)}\left(\frac{\|f_h(t) - f_{2\,h}(t)\|_p}{\|f_0\|_p}\right), \qquad 1\le p\le+\infty,$$

where f_h represents the approximation computed from a grid of order h. The numerical scheme is said to be k-th order if $\varepsilon_{2\,h} \leq C\,h^k$, for all $0 < h \ll 1$.



(1) (2)

FIGURE 2. The L^1 and L^∞ errors of the second order method (2.3) for different values of the Knudsen number $\varepsilon = 10^{-5}, \ldots, 1$.

In Figure 2, the L^1 and L^∞ errors of the second order method (2.3) are presented. They show a uniformly second order convergence rate (an estimation of the slope is 1.9) in space and time (the velocity discretization is spectrally accuracy in v thus does not contribute much to the errors). The time step is not constrained by the value of ε, showing a uniform stability in time.

4.2. **The Sod tube problem.** This test deals with the numerical solution to the $1d_x \times 2d_v$ Boltzmann equation for Maxwellian molecules ($\gamma = 0$). We present numerical simulations for one dimensional Riemann problem and compute an approximation for different Knudsen numbers, from rarefied regime to the fluid regime.

Here, the initial data corresponding to the Boltzmann equations are given by the Maxwellian distributions computed from the following macroscopic quantities

$$\begin{cases} (\rho_l, u_l, T_l) = (1, 0, 1)\,, & \text{if } 0 \leq x \leq 0.5\,, \\ \\ (\rho_r, u_r, T_r) = (0.125, 0, 0.25)\,, & \text{if } 0.5 < x \leq 1\,. \end{cases}$$

We perform several computations for $\varepsilon = 1, 10^{-1}, 10^{-2}$,...,$10^{-4}$. In Figure 3, we only show the results obtained in the kinetic regime (10^{-2}) using a spectral scheme for the discretization of the collision operator [22] (with $n_v = 32^2$ and a truncation of the velocity domain $v_{\max} = 7$) and second order explicit Runge-Kutta and second order method (2.3) for the time discretization with a time step $\Delta t = 0.005$ satisfying the CFL condition for the transport part (with $n_x = 100$). For such a value of ε, the problem is not stiff and this test is only performed to compare the accuracy of our second order scheme (2.3) with the classical (second order) Runge-Kutta method. We present several snapshots of the density, mean velocity, temperature and heat flux

$$\mathbb{Q}(t,x) := \frac{1}{\varepsilon} \int_{\mathbb{R}^{d_v}} (v - u_\varepsilon)\,|v - u_\varepsilon|^2\, f_\varepsilon(t,x,v) dv$$

at different time $t = 0.10$ and 0.20. Both results agree well with only $n_x = 100$ in the space domain and $n_v = 32$ for the velocity space. Thus, in the kinetic regime our second order method (2.3) gives the same accuracy as a second order fully explicit scheme without any additional computational effort.

Now, we investigate the cases of small values of ε for which an explicit scheme requires the time step to be of order $O(\varepsilon)$. In order to evaluate the accuracy of our method (2.3) in the Navier-Stokes regime (for small $\varepsilon \ll 1$ but not negligible), we compared the numerical solution for $\varepsilon = 10^{-3}$ with one obtained with a small time step $\Delta t = O(\varepsilon)$ (for which the computation is still feasible). Note that a direct comparison with the numerical solution to the compressible Navier-Stokes system (1.7) is difficult since the viscosity $\mu_\varepsilon = \mu(T_\varepsilon)$ and the thermal conductivity $\kappa_\varepsilon = \kappa(T_\varepsilon)$ are not explicitly known. Therefore, in Figure 4, we report the numerical results for $\varepsilon = 10^{-3}$ and propose a comparison

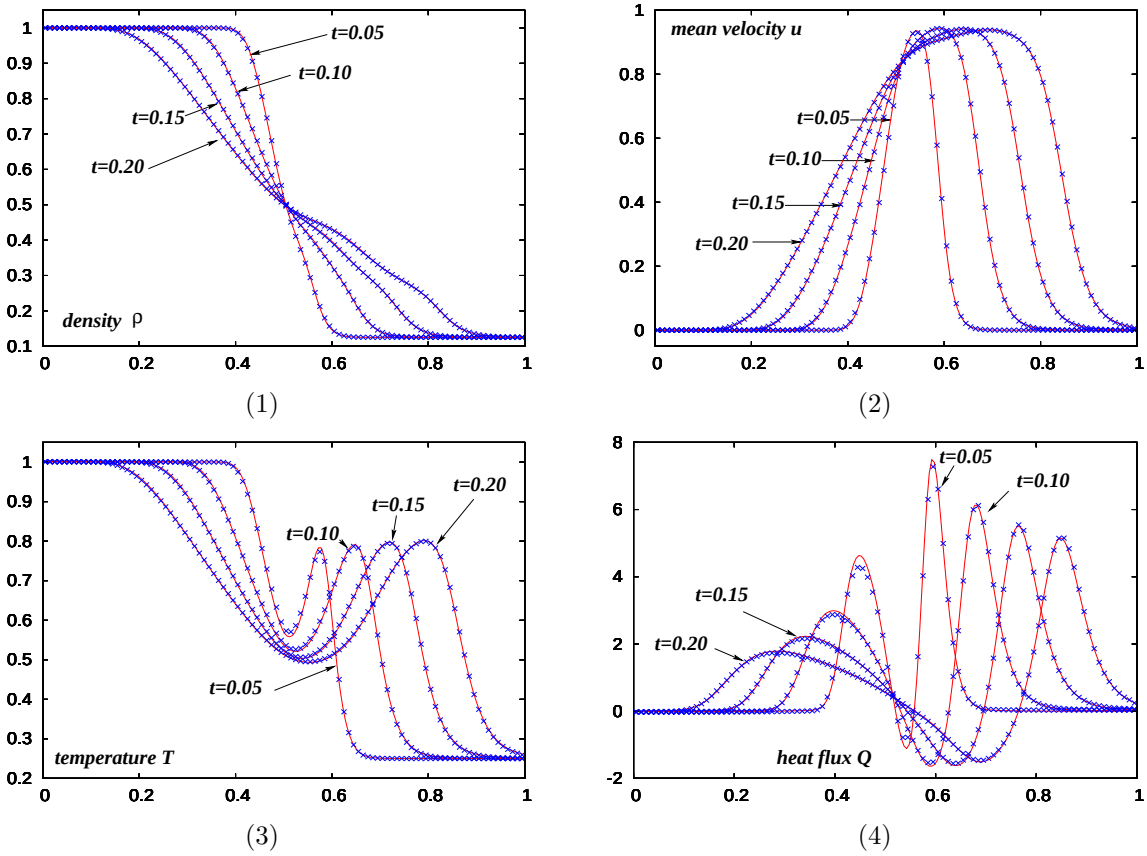


FIGURE 3. Sod tube problem ($\varepsilon = 10^{-2}$), dots (x) represent the numerical solution obtained with our second order method (2.3) and lines with the Runge-Kutta method: evolution of (1) the density ρ, (2) mean velocity u, (3) temperature T and (4) heat flux $\mathbb{Q}$ at time $t = 0.05$, 0.1, 0.15 and 0.2.

between the numerical solution obtained with the scheme (2.3) and the one obtained with a second order explicit Runge-Kutta method. In this case, the behavior of macroscopic quantities (density, mean velocity, temperature and heat flux) agree very well even it the time step is at least ten times larger with our method (2.2) or (2.3).

Finally in Figure 5, we compare the numerical solution of the Boltzmann equation (1.1) with the numerical solution to the compressible Navier-Stokes system derived from the BGK model since the viscosity and heat conductivity are in that case explicitly known [2]. To approximate the compressible Navier-Stokes system, we apply a second order Lax-Friedrich scheme using a large number of points ($n_x = 1000$) whereas we only used $n_x = 100$, and 200 points in space and $n_v^2 = 32^2$ points in velocity for the approximation of the kinetic equation (1.1). In this problem, the density, mean velocity and temperature are relatively close to the one obtained with the approximation of the Navier-Stokes system. Even the qualitative behavior of the heat flux agrees well with the heat flux corresponding to the compressible Navier-Stokes system $\kappa_\varepsilon \nabla_x T_\varepsilon$, with $\kappa_\varepsilon = \rho_\varepsilon T_\varepsilon$ (see Figure 5), yet some differences can be observed, which means that the use of BGK models to derive macroscopic models has a strong influence on the heat flux.

4.3. **A problem with mixing regimes.** Now we consider the Boltzmann equation (1.1) with the Knudsen number $\varepsilon > 0$ depending on the space variable in a wide range of mixing scales.

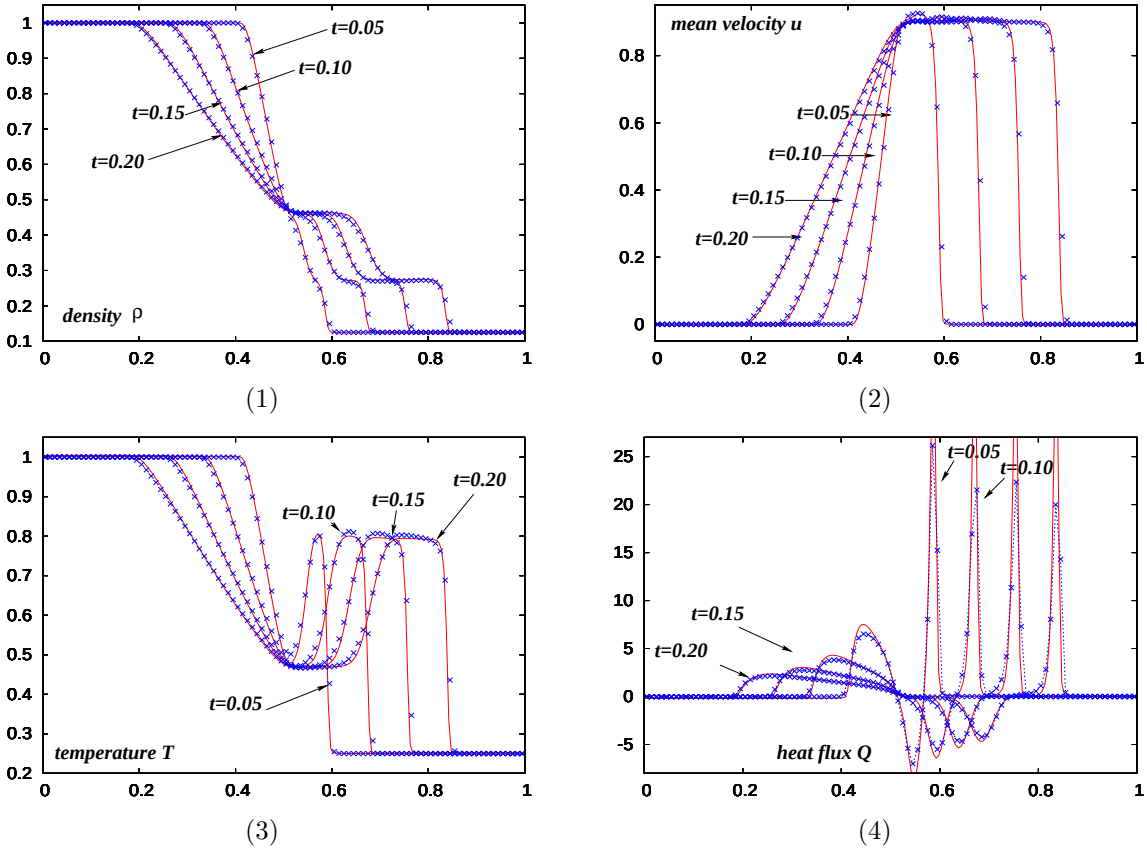
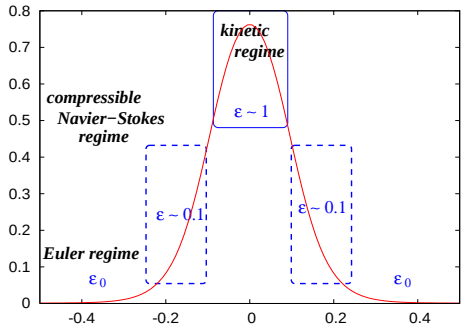


FIGURE 4. Sod tube problem ($\varepsilon = 10^{-3}$), dots (x) represent the numerical solution obtained with our second order method (2.3) and lines with the Runge-Kutta method: evolution of (1) the density ρ, (2) mean velocity u, (3) temperature T and (4) heat flux $\mathbb{Q}$ at time $t = 0.05$, 0.1, 0.15 and 0.2.

This kind of problem was already studied by several authors for the BGK model [17] or the radiative transfer equation [36]. In this problem, $\varepsilon : \mathbb{R} \mapsto \mathbb{R}^+$ is given by

$$\varepsilon(x) = \varepsilon_0 + \frac{1}{2}\left[\tanh(1 - 11\,x) + \tanh(1 + 11\,x)\right],$$

which varies smoothly from ε_0 to $O(1)$.



This numerical test is difficult because different scales are involved. It requires a good accuracy of the numerical scheme for all range of ε. In order to focus on the multi-scale nature we only consider periodic boundary conditions, even if the method has also been used with specular reflection in space. Furthermore, to increase the difficulty we consider an initial data which is far from the local equilibrium of the collision operator:

$$f_0(x,v) = \frac{\rho_0}{2}\left[\exp\left(-\frac{|v-u_0|^2}{T}\right) + \exp\left(-\frac{|v+u_0|^2}{T_0}\right)\right], \quad x \in [-L,L],\ v \in \mathbb{R}^2$$

with $u_0 = (3/4, -3/4)$,

$$\rho_0(x) = \frac{2+\sin(k\,x)}{2}, \quad T_0(x) = \frac{5+2\,\cos(k\,x)}{20}$$

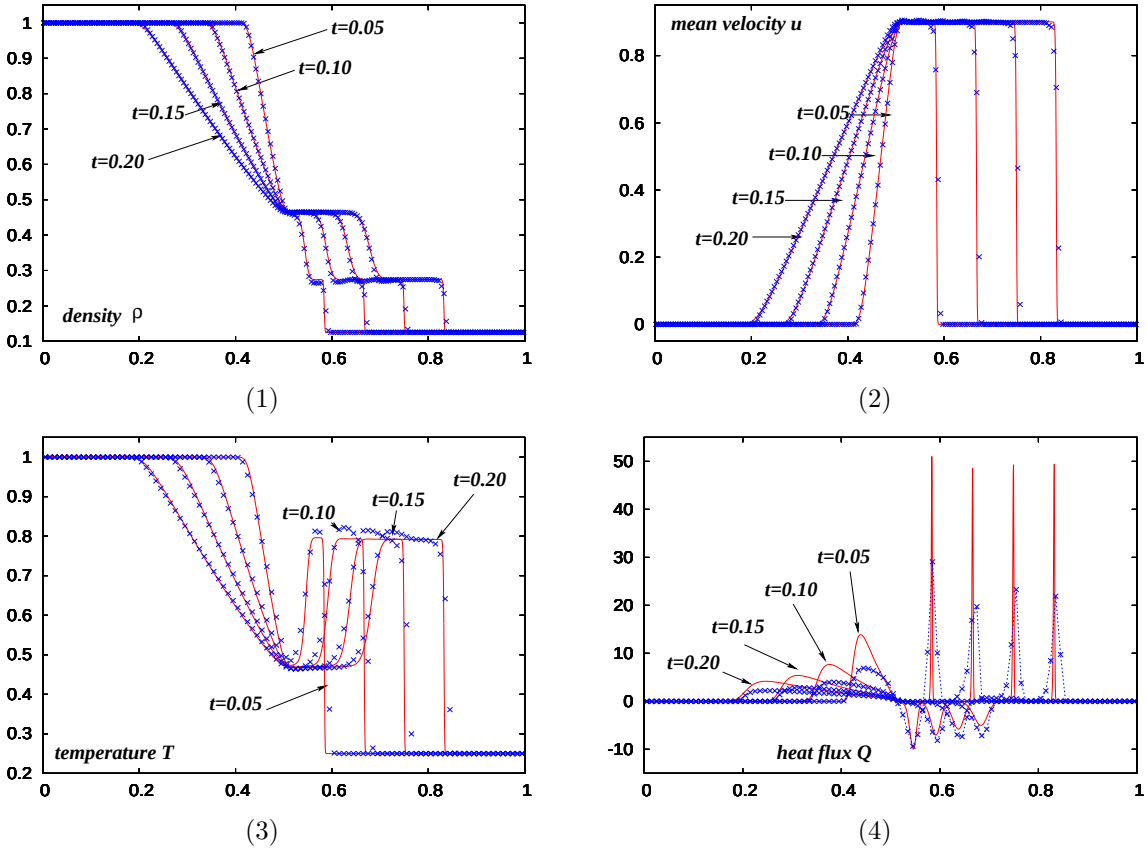


FIGURE 5. Sod tube problem ($\varepsilon = 10^{-4}$), comparison of the numerical solution to the Boltzmann equation with the our second order method (2.3) represented with dots (x) with the numerical solution to the compressible Navier-Stokes system (lines): evolution of (1) the density ρ, (2) mean velocity u, (3) temperature T and (4) heat flux $\mathbb{Q}$ at time $t = 0.05$, 0.1, 0.15 and 0.2.

where $k = \pi/L$ and $L = 1/2$.

Here we cannot compare the numerical solution with the one obtained by a macroscopic model. From the numerical simulations, we observe that the solution is smooth during a short time and some discontinuities are formed in the region where the Knudsen number ε is very small and then propagate into the physical domain.

On the one hand, we only take $\varepsilon_0 = 10^{-3}$ in order to propose a comparison of numerical solutions computed with a second order method using a time step $\Delta t = 0.001$ (such that the CFL condition for the transport part is satisfied) and the one by the second order explicit Runge-Kutta method with a smaller time step $\Delta t = 0.0001$) to get stability. The number of points in space is $n_x = 200$ and in velocity is $n_v^2 = 32^2$. Clearly, in Figure 6, the results are in good agreement even if our new method does not solve accurately small time scales when the solution is for from the local equilibrium. Moreover in Figure 7, we present numerical results with only $n_x = 50$ and $n_x = 200$, and $n_v^2 = 32^2$ to show the performance of the method with a small number of discretization points in space. With $n_x = 50$ points the qualitative behavior of the macroscopic quantities (ρ, u, T) is fairly good.

On the other hand, we have performed different numerical results when $\varepsilon_0 = 10^{-4}$, then the variations of ε starts from 10^{-4} to 1 in the space domain. In that case, the computational time of a fully explicit scheme would be more than one hundred times larger than the one required for the asymptotic preserving scheme (2.3). We observe that discontinuities appear on the density, mean velocity and temperature and then propagate accurately into the domain. The shock speed is roughly the same for the different numerical resolutions. Therefore, this method gives a very good compromise

between accuracy and stability for the different regimes. Numerical results are not plotted since they are relatively close to the ones presented in Figures 6 and 7.

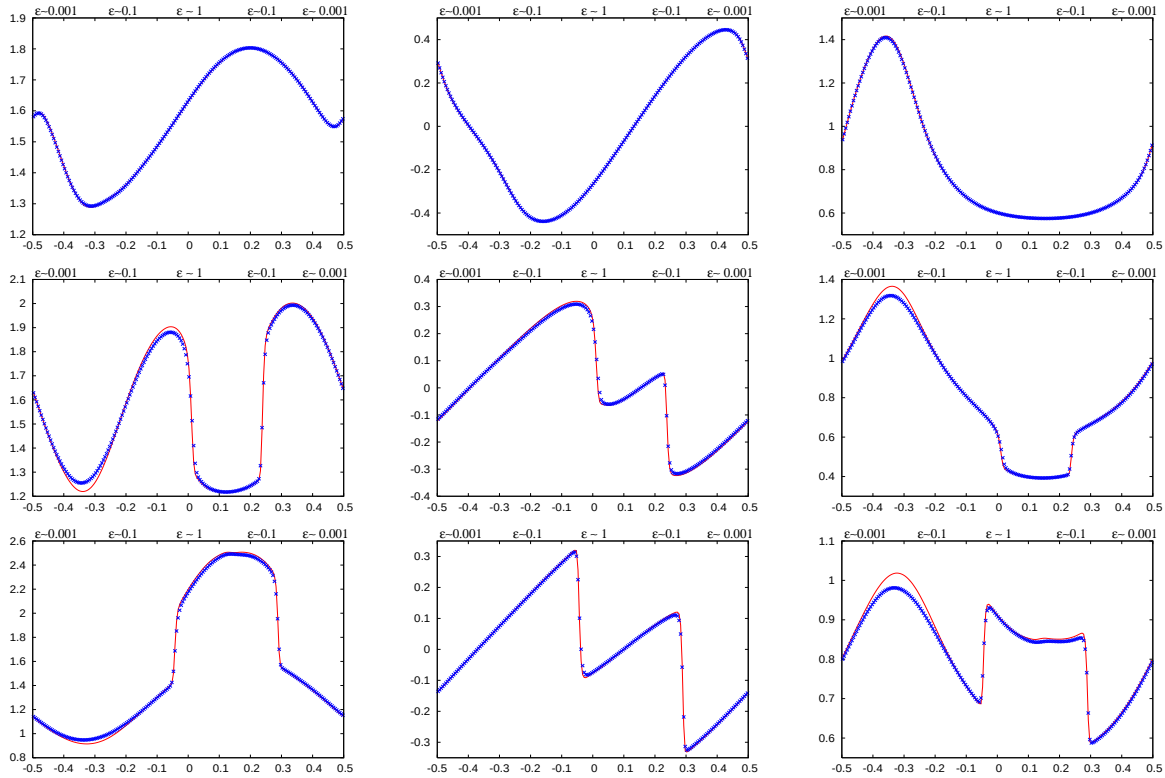


FIGURE 6. Mixing regime problem ($\varepsilon_0 = 10^{-3}$), comparison of the numerical solution to the Boltzmann equation with the second order method (2.3) represented with dots (x) with the numerical solution obtained with the explicit Runge-Kutta method using a small time step (line): evolution of (1) the density ρ, (2) mean velocity u, (3) temperature T at time $t = 0.25$, 0.5 and 0.75.

5. Other applications: numerical stability

In this section, we want to illustate the efficiency of the asymptotic preserving scheme to treat high order differential operators. We have already applied such a scheme for Willmore flow (fourth order differential operator [24, 42]). Here, we consider the flow of gas in a two dimensional porous medium with initial density $g_0(v) \geq 0$. The distribution function $g(t, v)$ then satisfies the nonlinear degenerate parabolic equation

$$\frac{\partial g}{\partial t} = \Delta_v g^m, \tag{5.1}$$

where $m > 1$ is a physical constant. Assuming that

$$\int_{\mathbb{R}^2} (1 + |v|^2)\, g_0(v) dv < +\infty,$$

J.A. Carrillo and G. Toscani [7] proved that $g(t, v)$ behaves asymptotically in a self-similar way like the Barenblatt-Pattle solution, as $t \to +\infty$. More precisely, it is easy to see that if we consider the change of variables

$$g(t, v) = \frac{1}{s(t)} f\left(\log(s(t)), \frac{v}{s(t)}\right), \tag{5.2}$$

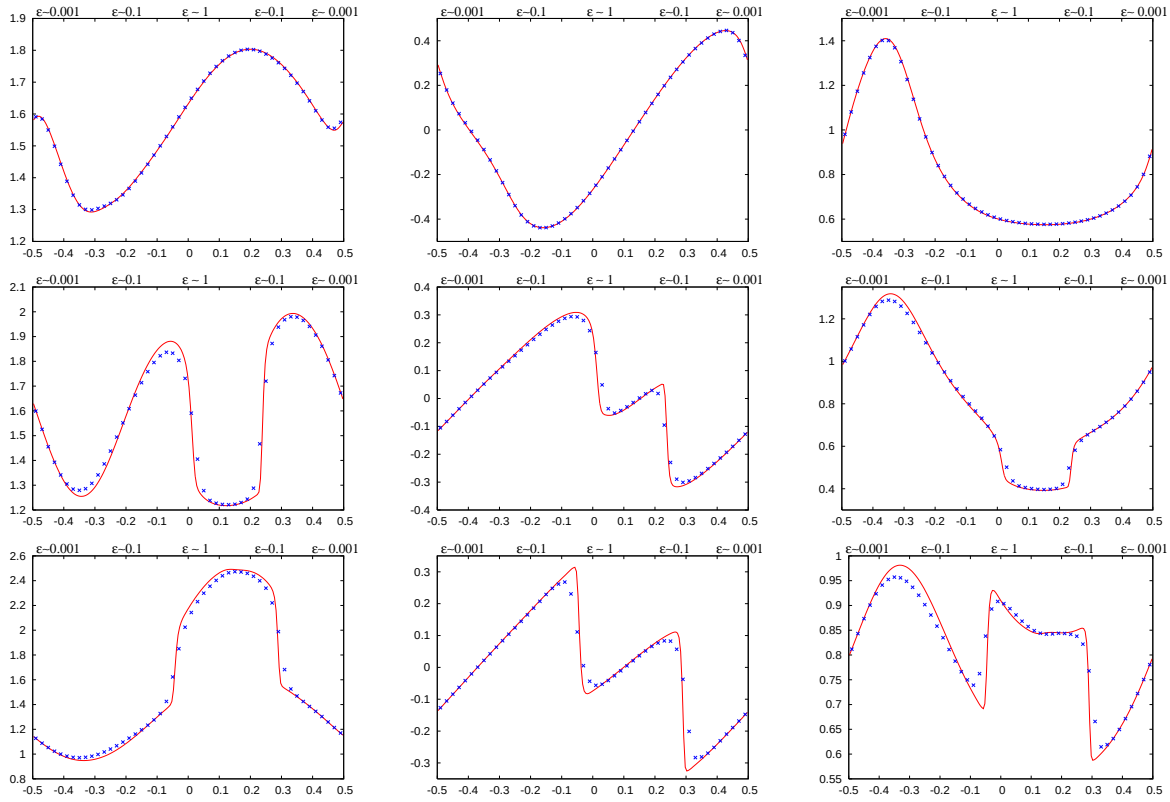


FIGURE 7. Mixing regime problem ($\varepsilon_0 = 10^{-3}$), comparison of the numerical solution to the Boltzmann equation obtained with the AP scheme (2.3) using $n_x = 50$ (dots x) and $n_x = 200$ points (line): evolution of (1) the density ρ, (2) mean velocity u, (3) temperature T at time $t = 0.25$, 0.5 and 0.75.

where $s(t) := \sqrt{1+2t}$, the new distribution function f is solution to

$$\frac{\partial f}{\partial t} = \nabla_v \cdot (v\, f \,+\, \nabla_v f^m)\,,$$

and converges to the Barenblatt-Pattle distribution

$$\mathcal{M}(v) = \left(C - \frac{m-1}{2\,m}\,|v|^2 \right)_+^{1/(m-1)},$$

where C is uniquely determined and depends on the initial mass g_0 but not on the "details" of the initial data.

Instead of working on (5.1) directly, we will study the asymptotic decay towards its equilibrium. The key argument on the proof of J.A. Carrillo and G. Toscani is the control of the entropy functional

$$H(f) = \int_{\mathbb{R}^2} \left[|v|^2\, f(t,v) + \frac{m}{m-1} f^m(t,v) \right] dv,$$

which satisfies

$$\frac{dH(f)}{dt} = -\int_{\mathbb{R}^2} f(t,v) \left| v + \frac{m}{m-1} \nabla f^{m-1} \right|^2 dv \le 0$$

or the control of the relative entropy $H(f|\mathcal{M}) = H(f) - H(\mathcal{M})$ with respect to the steady state $\mathcal{M}$.

Numerical discretization of this problem leads to the following difficulty : explicit schemes are constrained by a CFL condition $\Delta t \simeq \Delta v^2$ whereas implicit schemes require the numerical resolution of a nonlinear problem at each time step (with a local constraint on the time step). We refer to [9, 21] for a fully implicit approximation preserving steady states for nonlinear Fokker-Planck type equations.

Here we do not focus on the velocity discretization, but only want to apply our splitting operator technique to remove this severe constraint on the time step. Here the parameter ε does not represent a physical time scale but is only related to the velocity space discretization Δv. Therefore, we set $Q(f) = \nabla_v \cdot (v\, f \,+\, \nabla_v f^m)$ and $P(f) = \nabla Q(\mathcal{M})\,(f - \mathcal{M})$, which leads to the following decomposition

$$\frac{\partial f}{\partial t} \,=\, \underbrace{\Delta_v \left(f^m \,-\, m\, \mathcal{M}^{m-1}\, f\right)}_{\text{non stiff part}} \,+\, \underbrace{\nabla_v \cdot \left(v\, f \,+\, m\, \nabla_v \left(\mathcal{M}^{m-1}\, f\right)\right)}_{\text{stiff linear part}}.$$

Then we apply a simple IMEX scheme which only requires the numerical resolution of a linear system at each time step.

We choose $m = 3$ and a discontinuous initial datum far from the equilibrium

$$f_0(v) = \sum_{l\in\{1,2\}} \sum_{k\in\{0,\dots,n-1\}} \frac{1}{10}\, \mathbf{1}_{\mathcal{B}(0,r_0)}(v - v_{k,l})$$

where $n = 12$, $r_0 = 1/4$ and $v_{k,l} \,= l\, \mathrm{e}^{i\,\theta_k}$, with $\theta_k = \,2\,k\,\pi/n$, $k = 0, \dots, n-1$. We use a standard velocity discretization in the velocity space based on an upwind finite volume approximation for the transport term and a center difference for the diffusive part. We take $n_v^2 = 120^2$ in velocity and a time step $\Delta t = 0.02$ which is much larger than the time step satisfying a classical CFL condition for this problem $\Delta t \simeq O(\Delta v^2)$. The numerical scheme (2.2) is still stable and the numerical solution preserves nonegativity at each time step (see Figure 8)! For large time, the solution converges to an approximation of the steady state even if the present scheme is not exactly well-balanced (it does not preserve exactly the steady state). Moreover, to get a better idea on the behavior of the numerical solution, we plot the evolution of the entropy and its dissipation for different time steps. More surprisingly, the numerical entropy is decreasing and the dissipation converges towards zero when times goes to infinity.

6. Conclusion

We have proposed a new class of numerical schemes for physical problems with multiple time and spatial scales described by a still nonlinear source term. A prototype equation of this type is the Boltzmann equation for rarified gas. When the Knudsen number is small, the stiff collision term of the Boltzmann equation drives the density distribution to the local Maxwellian, thus the macroscopic quantities such as mass, velocity and temperature are be evolved according to fluid dynamic equations such as the Euler or Navier-Stokes equations. Asmptotic-preserving (AP) schemes for kinetic equations have been successful since they capture the fluid dynamic behavior even without numerically resolving the small Knudsen number. However, the AP schemes need to treat the stiff collision terms implicitly, thus it yeilds a complicated numerical algebraic problem due to the nonlinearity and nonlocality of the collision term. In this paper, we propose to augement the nonlinear Boltzmann collision operator by a much simpler BGK collision operator, and impose implicity only on the BGK operators which can be handled much more easily. We show that this method is AP in the Euler regime, and is also consistent to the Navier-Stokes approximations for suitably small time steps and mesh sizes. Numerical examples, including those with mixing scales and non-local-Maxwellian initial data, decomstrate the AP property as well as uniform convergence (in the Knudsen number) of this method.

This method can be extended to a wide class of PDEs (or ODEs) with stiff source terms that admit a stable and unique local equilibirum. We use the Fokker-Planck equation as an example to illustrate this point, and will pursue more applications in the future.

Acknowledgments. F. Filbet thanks Ph. Laurençot, M. Lemou, P. Degond and L. Pareschi for usefull discussions on the topic.

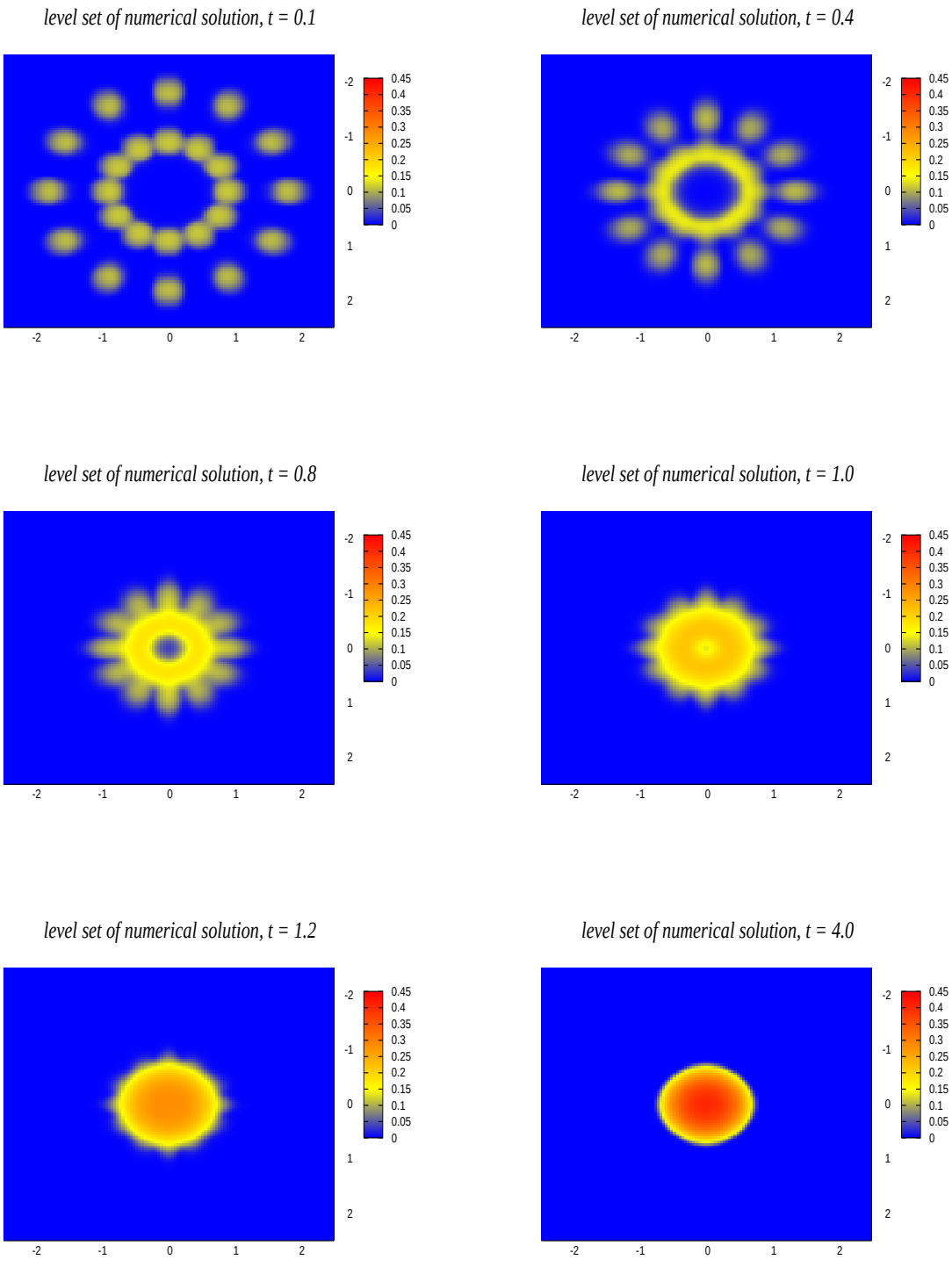


FIGURE 8. Nonlinear Fokker-Planck solution: convergence toward equilibrium (Barenblatt-Pattle distribution) obtained with the first order method (2.2) using $n_x = 100$ at time $t = 0.1$, 0.4, 0.8, 1.0, 1.2 and 4 with a large time step.

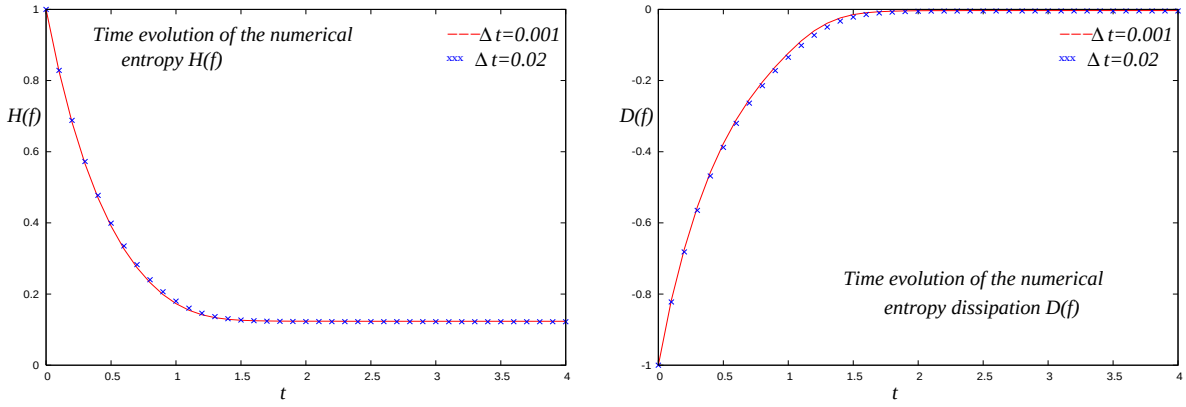


FIGURE 9. Nonlinear Fokker-Planck solution: convergence toward equilibrium (Barenblatt-Pattle distribution) obtained with the first order method (2.2) using $n_x = 100$ with $\Delta t = 0.02$ and 0.001.

REFERENCES

[1] C. Bardos; F. Golse and D. Levermore, Fluid dynamic limits of kinetic equations. I. Formal derivations. *J. Statist. Phys.* **63** (1991), 323–344.

[2] M. Bennoune; M. Lemou and L. Mieussens, Uniformly stable numerical schemes for the Boltzmann equation preserving the compressible NavierStokes asymptotics, *J. Comput. Phys.* **227** (2008), 3781–3803.

[3] P. L. Bhatnagar, E. P. Gross and K. Krook, A model for collision processes in gases, Phys. Rev. 94 (1954) 511-524

[4] J.-F. Bourgat, P. Le Tallec, B. Perthame, and Y. Qiu, Coupling Boltzmann and Euler equations without overlapping, in Domain decomposition methods in science and engineering (Como, 1992), 377–398, Contemp. Math. **157**, Amer. Math. Soc., Providence, RI, 1994.

[5] R. Caflish; S. Jin and G. Russo, Uniformly accurate schemes for hyperbolic systems with relaxation, *SIAM J. Numer. Anal.* **34** (1997) 246281.

[6] R.E. Caflisch and L. Pareschi, An implicit Monte Carlo method for rarefied gas dynamics I: The space homogeneous case, J. Computational Physics, 154, pp. 90-116, (1999).

[7] J. A. Carrillo and G. Toscani Asymptotic L^1-decay of solutions of the porous medium equation to self-similarity. *Indiana Univ. Math. J.* **49** (2000), pp. 113–142.

[8] C. Cercignani, The Boltzmann equation and its applications, Springer, 1998.

[9] C. Chainais-Hillairet and F. Filbet, Asymptotic behavior of a finite volume scheme for the transient drift-diffusion model, *IMA J. Num. Anal.* **27**, (2007) 689–716.

[10] G.Q. Chen, T.P. Liu and C.D. Levermore, Hyperbolic conservation laws with stiff relaxation terms and entropy. Comm. Pure Appl. Math. 47 (1994), no. 6, 787–830.

[11] F. Coquel and B. Perthame, Relaxation of energy and approximate Riemann solvers for general pressure laws in fluid dynamics. (English summary) SIAM J. Numer. Anal. 35 (1998), no. 6, 2223–2249

[12] F. Coron and B. Perthame, Numerical passage from kinetic to fluid equations, *SIAM J. Numer. Anal.* **28** (1991) pp. 26-42.

[13] P. Crispel; P. Degond, and M.-H. Vignal, An asymptotically preserving scheme for the two-fluid Euler-Poisson model in the quasi-neutral limit, *J. Comput. Phys.*, **223** (2007), pp. 208234.

[14] P. Degond; F. Deluzet and L. Navoret, An asymptotically stable Particle-in-Cell (PIC) scheme for collisionless plasma simulations near quasineutrality, *C. R. Acad. Sci. Paris Sér. I Math.*, **343** (2006), pp. 613618.

[15] P. Degond and S. Jin, A smooth transition model between kinetic and diffusion equations, SIAM J. Numer. Anal. 42 (6) (2005) 2671-2687

[16] P. Degond; S. Jin and J.-G. Liu, Mach-number uniform asymptotic-preserving gauge schemes for compressible flows. *Bull. Inst. Math. Acad. Sin. (N.S.)* **2** (2007), pp. 851–892.

[17] P. Degond; S. Jin and L. Mieussens, A smooth transition model between kinetic and hydrodynamic equations. *J. Comput. Phys.* **209** (2005), pp. 665–694.

[18] F. Filbet and L. Pareschi, A numerical method for the accurate solution of the Fokker-Planck-Landau equation in the non homogeneous case. *J. Comput. Phys.* **179**, (2002) pp. 1–26.

[19] F. Filbet and G. Russo, High order numerical methods for the space non-homogeneous Boltzmann equation. *J. Comput. Phys.* **186**, (2003) pp. 457–480.

[20] F. Filbet; L. Pareschi and G. Toscani, Accurate numerical methods for the collisional motion of (heated) granular flows. *J. Comput. Phys.* **202**, (2005) pp. 216–235.

[21] F. Filbet, A finite volume scheme for the Patlak-Keller-Segel chemotaxis model, *Numerische Mathematik*, **104** (2006) pp. 457–488.

[22] F. Filbet; C. Mouhot and L. Pareschi, Solving the Boltzmann equation in $N \log_2 N$. *SIAM J. Sci. Comput.* **28**, (2006) pp. 1029–1053

[23] F. Filbet, An asymptotically stable scheme for diffusive coagulation-fragmentation models, Comm. Math. Sciences, **6**, (2008) pp. 257–280.
[24] F. Filbet and C.W. Shu, work in progress
[25] E. Gabetta; L. Pareschi, and G. Toscani, Relaxation schemes for nonlinear kinetic equations, *SIAM J. Numer. Anal.* **34** (1997), 2168–2194
[26] C.W. Gear, Numerical Initial Value Problems in Ordinary Differential Equations, Prentice Hall, 1971.
[27] F. Golse, S. Jin and C.D. Levermore, The Convergence of Numerical Transfer Schemes in Diffusive Regimes I: The Discrete-Ordinate Method, SIAM J. Num. Anal. 36 (1999) 1333-1369
[28] L. Gosse and G. Toscani, An asymptotic-preserving well-balanced scheme for the hyperbolic heat equations, C.R. Math. Acad. Sci. Paris 334 (2002) 337-342
[29] M. Günther, P. Le Tallec, J.-P. Perlat, and J. Struckmeier, Numerical modeling of gas flows in the transition between rarefied and continuum regimes. Numerical flow simulation I, (Marseille, 1997), 222–241, Notes Numer. Fluid Mech., 66, Vieweg, Braunschweig, 1998.
[30] J. Haack, S. Jin and J.-G. Liu, An all-speed asymptotic-preserving schemes for compressible flows, in preparation.
[31] S. Jin, Effcient asymptotic-preserving (AP) schemes for some multiscale kinetic equations, *SIAM J. Sci. Comput.* **21** (1999) pp. 441-454,
[32] S. Jin Runge-Kutta Methods for Hyperbolic Conservation Laws with Stiff Relaxation Terms, J. Computational Physics, **122** (1995), 51-67.
[33] S. Jin and C.D. Levermore, Numerical schemes for hyperbolic conservation laws with stiff relaxation terms. J. Comput. Phys. 126 (1996), no. 2, 449–467.
[34] S. Jin and L. Pareschi, Discretization of the multiscale semiconductor Boltzmann equation by diffusive relaxation schemes, *J. Comput. Phys.* **161** (2000) 312-330.
[35] S. Jin, L. Pareschi and G. Toscani, Diffusive Relaxation Schemes for Discrete-Velocity Kinetic Equations, SIAM J. Num. Anal. 35 (1998) 2405-2439
[36] S. Jin; L. Pareschi and G. Toscani, Uniformly accurate diffusive relaxation schemes for multiscale transport equations. *SIAM J. Numer. Anal.* **38** (2000), 913–936
[37] A. Klar, An asymptotic-induced scheme for nonstationary transport equations in the diffusive limit, *SIAM J. Numer. Anal.* **35** (1998) 10731094,
[38] A. Klar, An asymptotic preserving numerical scheme for kinetic equations in the low Mach number limit, *SIAM J. Numer. Anal.* **36** (1999) 15071527.
[39] A. Klar, H. Neunzert, and J. Struckmeier, Transition from kinetic theory to macroscopic fluid equations: a problem for domain decomposition and a source for new algorithm, Transp. Theory and Stat. Phys. 29 (2000) 93-106
[40] M. Lemou and L. Mieussens, A new asymptotic preserving scheme based on micro-macro formulation for linear kinetic equations in the diffusion limit. SIAM J. Sci. Comput. 31 (2008) no. 1, 334-368.
[41] L. Pareschi and G. Russo, Time relaxed Monte Carlo methods for the Boltzmann equation, *SIAM J. Sci. Comput.* **23** (2001) 1253-1273,
[42] P. Smereka Semi-implicit level set methods for curvature and surface diffusion motion. Special issue in honor of the sixtieth birthday of Stanley Osher. *J. Sci. Comput.* **19** (2003) pp. 439–456.
[43] P. Le Tallec, and F. Mallinger, Coupling Boltzmann and Navier-Stoke s equations by half fluxes, J. Comput. Phys. 136 (1997) 51-67
[44] H.C. Yee, A Class of High-Resolution Explicit and Implicit Shock-Capturing Methods, Von Karman Institute for Fluid Dynamics Lecture Series, 1989

Francis Filbet

Université de Lyon,
Université Lyon I, CNRS
UMR 5208, Institut Camille Jordan
43, Boulevard du 11 Novembre 1918
69622 Villeurbanne cedex, FRANCE

e-mail: filbet@math.univ-lyon1.fr

Shi Jin

Department of Mathematics
University of Wisconsin
Madison, WI 53706, USA

e-mail: jin@math.wisc.edu